Representation of Many-Body Interactions with Pairwise Descriptors on Machine Learning Interatomic Potential

Akira Takahashi¹, Atsuto Seko¹,²,³,⁴, Isao Tanaka¹,²,⁴,⁵

¹Department of Materials Science and Engineering, Kyoto University, Kyoto, 606-8501, Japan
²Center for Elements Strategy Initiative for Structure Materials (ESISM), Kyoto University, Kyoto 606-8501, Japan
³JST, PRESTO, Kawaguchi 332-0012, Japan
⁴Center for Materials Research by Information Integration, National Institute for Materials Science, Tsukuba 305-0047, Japan
⁵Nanostructures Research Laboratory, Japan Fine Ceramics Center, Nagoya, 456-8587, Japan

Interatomic potentials have been widely used to perform a large-scale molecular dynamics simulation. The reliability and computational cost of the simulation crucially depend on the quality of the interatomic potential. Recently, methods to construct accurate interatomic potentials using a combination of a large set of density functional theory (DFT) calculations and regression methods were proposed [1-3]. This is called machine learning interatomic potential (MLIP). In MLIP, the relationship between the energy and a set of descriptors representing the crystal structure is estimated from a DFT data set. However, its physical interpretation or the relationship between MLIP and existing conventional interatomic potentials are still lacking.

In this study, we provide the relationship between linearized MLIP and conventional "pair functional interatomic potential", such as embedded atom model (EAM). The linearization of MLIP makes it easy to give some interpretation. At the same time, we generalize the linearized MLIP proposed in the previous paper [3] by introducing cross terms of different descriptors, which enables us to take many-body interactions into account more accurately. Finally, we demonstrate applications of the generalized framework to 31 non-magnetic metals to indicate its flexibility and generality.

For each metal, a set of DFT calculations is performed for 2700 structures. Interatomic potentials are then constructed using the DFT data and the linear ridge regression. The prediction error of the interatomic potential is measured by the root-mean square (RMS) difference between DFT and predicted energies by the interatomic potential. Figure 1 shows the dependence of the prediction error on the number of descriptors for (a) Mg and (b) Al. Since the computational cost of the cross terms is negligible, the number of descriptors determines the computational cost. As can be seen in Fig. 1, the introduction of the cross terms of descriptors decreases both the number of descriptors required for the convergence of the prediction error and the minimum value of the prediction error significantly. In other words, the introduction of the cross terms significantly improves both the accuracy and computational performance of the interatomic potential. Most of the other metals show the same tendency.

Figure 2 shows the prediction errors of the optimal interatomic potentials for the 31 metals. In most of the metals, the prediction errors decrease by considering the cross terms. Significant improvement of the prediction errors can be seen in Be, Al, Ga, In, and transition metals. It implies that many-body interactions are dominant in these elements representing covalent bonds. Nevertheless, the prediction errors are still more...
than 10 meV/atom in transition metals such as Hf, W, and Cr. This may be ascribed to the fact that only pairwise-distance descriptors and their cross terms are considered in this study. In addition to these descriptors, introduction of bond-angular dependent descriptors may improve the quality of the interatomic potentials for transition metals.

Acknowledgement
AT was supported by JSPS KAKENHI Grant Numbers (15J07315) and a Grant-in-Aid for Scientific Research on Innovative Areas “Nano Informatics” (Grant No. 25106005). AS was supported by PRESTO, JST. AS and IT were also supported by ”Materials research by Information Integration” Initiative (MI2I) from Japan Science and Technology Agency.

References

FIG. 1. Dependence of the prediction error of interatomic potential on the number of descriptors for (a) Mg and (b) Al. Red open circles and blue closed squares show the prediction errors of interatomic potentials constructed with cross terms of descriptors and without cross terms of descriptors, respectively.

FIG. 2. Prediction error of the optimal interatomic potential for 31 metals. Closed and open bars show the prediction errors of interatomic potentials constructed with the cross terms of descriptors and without the cross terms of descriptors, respectively.